

DOCUMENT RESUME

ED 220 788

CS 006 633

AUTHOR Baten, Lut
TITLE Text Comprehension: The Parameters of Difficulty in Narrative and Expository Prose Texts: A Redefinition of Readability.
PUB DATE [81]
NOTE 29p.; Research prepared at the University of Illinois. Table may not reproduce well due to small print.

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Discourse Analysis; *Models; Narration; *Readability; *Reading Comprehension; *Reading Processes; Reading Research; Secondary Education; Structural Analysis (Linguistics)
IDENTIFIERS Expository Text; *Prose Learning; *Text Structure

ABSTRACT

A study was conducted to define the linguistic and discourse parameters of text difficulty from the point of view of both the reader and the text in order to redefine readability and to provide an operational way of explaining processing difficulties of the near-mature reader. Subjects were 14-year-old students in the United States and the United Kingdom who read eight well-written short texts (maximum 280 words) as part of the IAEEA (International Association for the Evaluation of Educational Achievement) test on reading. On the basis of micro-structural analysis (of syntax, speed of closure, lexical analysis, given and new information structure, and lexically cohesive elements per text) and macro-structural analysis (of narrational and structural analysis and propositional analysis), a theoretical model of parameters indicating the degree of difficulty of discourse processing was formed and the general hypothesis that macro-structural and micro-structural parameters both influence comprehension (with higher level parameters not necessarily having more impact on comprehension) was confirmed. Application of the model to other languages further confirmed its validity. Among the implications of these findings for teachers is the need to recognize that comprehension is a process and not just a product and that the reader is central to this process. (JL)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

**TEXT COMPREHENSION: THE PARAMETERS OF DIFFICULTY
IN NARRATIVE AND EXPOSITORY PROSE TEXTS:
A REDEFINITION OF READABILITY**

By Lut Baten.

Lut Baten

1. Purpose and Problems of the Study

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

The purpose of this study was to define the linguistic and discourse parameters of text difficulty from the point of view of both the reader and the text, in order to redefine readability and to provide an operational way of explaining processing difficulties of the near mature reader. Thus the study was designed as a theoretical investigation aiming at practical appli-
cations in the classroom. In carrying out this research task, the following general principles were followed:

- a. It is necessary to analyze and define what it is that makes a text hard or easy on the linguistic and the discourse levels.
- b. Such a definition should be made from the point of view of how the reader processes a text. (Therefore, the purpose is to highlight the process itself, and not the product of comprehension.
- c. It is not possible to rely entirely on traditional readability research in which objectivity, in K. Goodman's words (Goodman, 1977), is often construed to mean that only directly observable aspects of things are legitimate concerns. It is more useful to search for a theory that would explain the processing difficulties of readers reading authentic texts.

To quote Wardrop (1977):

If it is true that people read to obtain meaning, and that the purpose of instructional reading is to enable the learner to obtain meaning, then research in reading should be oriented toward understanding how meaning is acquired from printed symbols, after which we can deal with the instructional technology with something stronger than an ad hoc approach.

2. Theoretical Basis of the Study and of the Developed Model

The study is multi-disciplinary in approach. In elucidating the problem of how readers process a text, the investigation draws on recent insights in cognitive psychology, more precisely schema theory as developed originally by Bartlett (1932) and the learning theory, developed by Gagné and by Van Parreren. The study mainly builds on recent research in discourse processing, however, specifically that of Kintsch, Van Dijk, Meyer, Perfetti and Lesgold, Pearson, and Spiro.

Language and text are studied from the point of view of linguistics, psycholinguistics and discourse analysis. The study relies mainly on work done by Quirk and Greenbaum, by Halliday and Hasan, by Osgood, Chafe, and Clark and Clark. For discourse analysis, the study makes use of the work done by P. Fries, G. Brown, Fishman, and Widdowson. Some principles of text-linguistics and discourse processing like those of Genette, Thorndike, and Kieras are sources of the text structure analysis. The lexical analysis owes much to the work of Engels, and the peda-linguistic approach to Engels and Ingram.

3. Method of Analysis

The study started with data from the IEA (International Association for the Evaluation of Educational Achievement) test on reading, administered in the early seventies. These data were collected in order to compare group performance and answer questions about the educational system. Instead of looking at the total scores, however, we were interested in a detailed task analysis. The purpose was to contribute to a theory of text comprehension by problematizing the input presented to the students. For this purpose we analyzed the passages from linguistic and cognitive points of view. On the basis of these analyses we were able to derive hypothetical parameters of discourse difficulty. These hypotheses were built into a model, which was then tested against the already available passage comprehension scores. In short, the procedure was to build a model on a multi-disciplinary analysis and then to test that model using available actuarial records.

The following method will be used in the reporting of the results of the study. In the first stage, the method of analysis used to get a measure of each of the parameters of the model is described as well as the extent to which the parameter measures appear to discriminate between a group of easier versus more difficult texts. In the second stage the model is tested more formally by examining the intercorrelations between the various parameters and their correlations with the empirical text passage scores. Used as a point of reference to ascertain empirical comprehension difficulty, were the comprehension scores computed on the basis of discrete point testing and the teachers' reference computed on the basis of teachers' judgements of the difficulty of every text after they had taught the texts to their classes and had the readers

process the texts the way they were used to in their classrooms. This judgement of text difficulty has the advantage that it is made at the time of the actual processing, by someone who knows the students and not after the processing is done.

On the basis of both judgements, one can distinguish between an easy and a hard text group. The easy group contains Paracutin, Paper, Ernenek, Desert Siege, whereas the hard texts are Fez, Plastic Shoes, Camels, Scientific Method. The starting point of the analysis is the organizational and functional system of language as it is realized in these eight texts. Different aspects of text, on a micro- and a macro-structural level, are analyzed quantitatively in relation to the performance of the readers. Four main parts can be distinguished: lexicon (including the information structure of given and new, defined and undefined concepts), syntax (including speed of closure), text structure (including cohesion and theme-rheme organization), and propositional analysis (both on the micro- and macro-structural level).

The analyses fall into two main bodies: micro-structural with lexical and syntactic analysis, and macro-structural with text-structure and propositional analysis. The analysis of coherence forms a transition between the two.

4. Analysis of Text Characteristics

The selected texts were read by fourteen-year-old students in the U.S.A. and the U.K. (Thorndike, 1973; Purves et al, 1981, also includes the text passages). The eight different texts were short (maximum 280 words), well-written and different in both style and content. Four were narrative (of which two were also descriptive), and four were expository.

4.1. Micro-structural Analysis

4.1.1. Syntax

Sentence structure, both sentence length and sentence complexity, was studied, based on Quirk and Greenbaum (1972), and also on Hunt's (1965) notion of T-units. The number of subordinate clauses preceding the main clause discriminated between hard and easy texts (see Table 1) in that if more subordinate clauses precede the finite verb of the main clause, the comprehension difficulty of the discourse is increased. The number of subordinate clauses to the right of the verb apparently bear no consequences for understanding. The

number of T-units per number of sentences had no impact on text difficulty either; neither had sentence length, as such. These results confirmed earlier findings by Simensen (1980) from a syntactic analysis of Norwegian texts and processing difficulty.

4.1.2. Speed of Closure

In relation to syntax, the semantic-syntactic units within the sentences of the passages were studied. Central in this is the primary mental ability called 'first closure factor' (Thurstone, 1944). This factor is defined as the ability to perceive an apparently disorganized or unrelated group of parts as a meaningful whole, i.e., the capacity to construct a whole picture from incomplete or limited material (Thurstone, 1966). This concept is also referred to as the "chunk model" because it claims that comprehension consists of synthesizing atomistic propositions into larger conceptual or semantic units rather than analyzing complex units into atomistic propositions (Pearson, 1973).

We hypothesized that a distinction between the hard and the easy text groups could be made on this basis. Therefore, we analyzed the "tonic breaks" readers insert in reading passages in order to make meaning. We understand "tonic breaks" in Halliday's (1967) sense: a unit in intonation with a certain intonation pattern and a stop at the end. For this analysis, we set up an experiment in which we had native speakers of English read the texts, unprepared, aloud, as if they were reading them silently. Their reading was recorded with the subjects alone in a small studio. Then, the subjects were asked to read the same texts aloud to an audience, the way they would for their students, in order to have them understand the texts as well as possible. In the first reading, the semantic units were longer in the hard than in the easy texts. In a second, edited version, however, they became much shorter in the hard passages and remained the same in the easy ones. This supports the hypothesis that comprehension involved chunking. In the more difficult texts smaller semantic units, which fit into a larger whole by means of their intonation, were indicated, whereas in the easy passages, no editing seemed to be necessary (see Table 1).

4.1.3. Lexical Analysis

In this analysis, the dominant measures of word difficulty which were studied are word frequency, word familiarity, and defining words.

A. Word frequency

We consulted the following word frequency lists:

- a. Kučera, H., and W. N. Francis (1967) Computational Analysis of Present Day American English.
- b. Engels, L. K., D. Geens and W. Martin (1975) Analysis of Present Day English Theatrical Language (1966 - 1972).
- c. Engles, L. K., B. Van Beckhoven, T. Leenders, I. Brasseur (1981) L.E.T. Vocabulary List, Leuven English Teaching Vocabulary List Based on Objective Frequency Combined with Subjective Word Selection, Acco, K.U.L. Department of Linguistics, section of Applied Linguistics.

The texts were matched with the three lists, but also counted were "outsiders," words occurring in the text, but not in these lists. Engels (1968) pointed out their importance by showing that contrary to the general belief that a list of 3000 words would cover 95% of the language (and thus enable a person to speak and to understand a foreign language by assimilating those words), only about 87% could be understood when subjects were presented random texts. Infrequent words are important because when topics increase, the number of outsiders increases and frequent words are often low in information content. It is therefore possible to understand 80% of the words of a text, and yet not to understand the "new" information of that text.

B. Word familiarity

The notion of familiarity in frequency counts has been questioned because of its subjective aspect (to one person, a word may sound familiar, but not to any one else), and its definition (What is "knowing" a word or one word meaning of a word?). J. C. Richards (1974) overcame these objections by compiling a familiarity count which is based upon the subjective impression of 1000 subjects, with respect to concrete words.

C. Coverage

In West's General Service List (1977) a criterion called "cover" is used. The concept "cover" accounts for the omission of a frequent word, if its meaning is already covered by another word or expression. "For the time being," for example, can be covered by "for the present." We could make use

of this list together with the compilation taken from Longman's Dictionary of Contemporary English (1978).

Both lists enabled us to check whether the vocabulary of the texts goes back to basic English and to a language used to explain concepts in a dictionary (definition vocabulary). Especially in expository prose, in which explanations should be as well comprehended as possible, these words are important.

Finally, and in relation to the three measures of word difficulty above, word length was studied, following Flesch (1948):

Matching the texts with the above lists brought the following results (see Table 1):

- a. there were more outsiders in the hard text group;
- b. the outsiders in the hard text group were morphologically more complex;
- c. there are more words appearing in the familiarity list in the easy text, group than in the hard text group;
- d. there are more words, drawn from coverage vocabulary in the easy texts than in the hard texts.

4.1.4. Given and New Information Structure

Here, we tried to elaborate the idea that the number of concepts determines the readability of a discourse. By means of an analysis of the given and new information structures rendered in discourse, we could distinguish between the load on the text processing activity per discourse. We relied on Mathesius (1939), Halliday and Hasan (1976), Chafe (1974), MacWhinney and Bates (1978), and Clark and Clark (1977) for the definition and the operationalization of the "given" and "new" information in discourse, both in terms of the formal-functional language organization and semantics. If a concept appeared for the first time in a text, it was labeled "new". If it was accompanied by an indefinite article or a dummy, it was labeled "indefinite". "Given" was assigned to any noun which was synonymous, repetitive or used as a collocation with any other noun used earlier in the text.

Between the hard and the easy text groups, there was no difference as to the number of concepts used. Moreover, contrary to the hypothesis, we found (see Table 1) that more "given" concepts occur in the hard texts and "new" in the easy texts. The latter, however, were annotated as "defined" or

Table 1. Summary of Measures of Studied Parameters of Text Characteristics.

Character- istic of Text	1	2	3		4		5		6		7				T- Units	Scenery		Summary		Pause	
			a	b	a	b	a	b	a	b	DG	UG	DN	UN		a	b	a	b	a	b
Paracutin	-	1	9.6	7.4	1	0.38	18	6.92	31	28	6.1	1.1	5.3	4.6	29	0	0	27	93	2	6
Ernenek	1.10	6	10.1	7.4	-	0.00	20	7.29	28	22	3.6	0.7	10.5	3.6	15	0	0	10	66	5	33
Paper	2.00	1	11.1	8.1	-	0.00	9	4.78	34	30	2.6	4.2	5.3	8.5	18	4	22	16	88	0	0
Desert Siege	-	5	10.6	7.3	-	0.00	12	7.01	33	28	5.2	0.5	11.7	6.4	16	2	12	14	87	0	0
Scientific Method	2.70	3	9.5	4.4	2	0.84	4	1.68	25	21	3.3	6.3	6.7	10.1	13	0	0	8	61	5	30
Camels	1.31	1	13.0	7.0	-	0.00	7	2.65	28	22	4.5	7.2	3.7	3.4	17	0	0	7	46	10	66
Plastic Shoes	0.83	-	12.5	6.0	-	0.00	5	2.46	34	30	2.9	0.9	0.5	3.9	12	0	0	2	22	11	122
Fez	6.14	10	8.6	5.6	6	3.14	3	1.57	30	26	2.0	1.5	6.8	8.9	7	2	26	5	71	3	42

Column

- 1 Total number of subordinate clauses to the left of the finite verb of the main clause.
- 2 Word frequency: outsiders
- 3 Speed of closure: number of words/intonation units in (a) version 1, (b) version 2.
- 4 Number of morphologically complex outsiders: (a) absolute number, (b) per 100 words.
- 5 Number of words appearing in Richard's list of familiarity: (a) absolute number, (b) per 100 words.
- 6 Number of words (in percentage) appearing in (a) the Longman vocabulary, and (b) West's list of coverage vocabulary.
- 7 Distribution of Given and New per 100 words: DG=defined/given, UG=undefined/given, DN=defined/new, UN=undefined/new.
- 8 Text duration: (a) absolute number, (b) per 100 T-units.

"identifiable", rather than as "undefined" or unidentifiable". We suggest that this "identifiable/new" indication would refer to an existing empathy between writer and reader. The writer knows what is identifiable, but at the same time new in his message for his audience. On the other hand, in the hard text group, more undefined, given concepts occur. Two conclusions can be drawn:

- a. it is not necessarily true that more "given" concepts occur in theme position;
- b. "defined" or "undefined" should be studied in connection with "given" and "new".

In summary, the analysis of the information focus in a text reveals the relationship between the formal-functional organization of language and the empathy of writer and reader. To further analyze this relationship, we have to focus on the conceptual fields themselves, however.

4.1.5. Lexically Cohesive Elements Per Text

Cohesion was studied in the sense of Halliday and Hasan (1976), who see it as a part of "coherence." Fishman (1977) found that "recoverable structural words" (anaphora, cataphora, etc.) did not contribute distinctively to comprehension difficulties (in relation to the cohesive aspect of the text). Therefore, we restricted ourselves to the lexical chains per passage and investigated whether they interacted in some consistent way with the texture of the texts.

Lexical chains were decided according to the following criteria:

- a. words having almost the same meaning;
- b. words semantically related (e.g., on a whole-part relationship or as hyponyms, etc.);
- c. words representing two poles;
- d. words calling forth other words, including associations.

These formed the lexical part of cohesion.

The grammatical aspect of cohesion is found in the relationship between cohesion and the texture of a text, more precisely in its theme-rheme organization. Texture of a text has to be understood in Halliday's sense. In general one can say that basic to texture are "theme" and "rheme." Theme is the point of departure of a text. In rheme-position, the main part of the message, the

"climatic feature" is found. In fact, this distinction of "theme" and "rheme" extends the limits of the sentence. We worked with the formal definition of theme and rheme, on a sentence level, to define the pattern of theme and rheme throughout a text.

Our basic question was: How strongly related are coherence, at the cohesion level, and comprehension? We adopted Halliday's (1980) condition that the cohesive ties in a text are not by themselves a guarantee of a coherent texture but that the resources have to be organized and deployed in patterned ways. On this basis, we formulated the hypothesis that, if a text displays coherence on the basis of cohesive ties (1) which are in cohesive harmony, and (2) which are consistently distributed in a theme-rheme distribution throughout the text, then such a text is more easily comprehended than a text in which there is a lower proportion of interaction of cohesive harmony in a theme-rheme display. The texts were then analyzed in this three-fold way:

- a. cohesive ties and their cohesive harmony;
- b. theme-rheme display;
- c. the interaction between the two.

The results showed that it is not the mere number of cohesive elements used in discourse that play a decisive role in the coherence and consequently in the processing of discourse, but their organization. The following prerequisites are proposed for a smooth processing:

- a. cohesive harmony (i.e., a consistent network of lexical chains);
- b. a consistent distribution of this harmony over theme-rheme display in the discourse (i.e., the same lexical fields in the same position, rheme or theme);
- c. congruence of the above with the overall text structure (i.e., in compliance with the expectations of a reader as to the linear and particular organization of discourse-structure, e.g., narrational, deductive or inductive development).

4.2. Macro-structural Analysis

4.2.1. Narrational and Structural Analysis

A three-fold comparison of the readers' abstract expectancy patterns of text organization was made for every passage: text duration, text order, and logical order. For text duration, we relied on the French structuralist school

of literary analysis (Genette, 1972; Fløttum, 1980). Within text duration, we distinguished between "summary," i.e., only the main actions are told (the time of telling is shorter than the action time); "pause," i.e., longer text time than action time (e.g., reflections); "scene," i.e., the time of telling is about the same as the time of the actual action.

For the analysis of text order, we distinguished between stories and expository prose texts. For the former we adopted Thorndyke's (1976) model of story structure, but we also applied Kieras' (1978) model for "framing" on the micro-structural level. The non-narrative texts were analyzed by means of Mountford's and Widdowson's models of expository prose (Mountford, 1975; Widdowson, 1979).

The results showed that text duration is a parameter of easy and hard texts. The easy texts appeared to contain much more "summary" whereas the hard texts contain much more "pause" (Table 1).

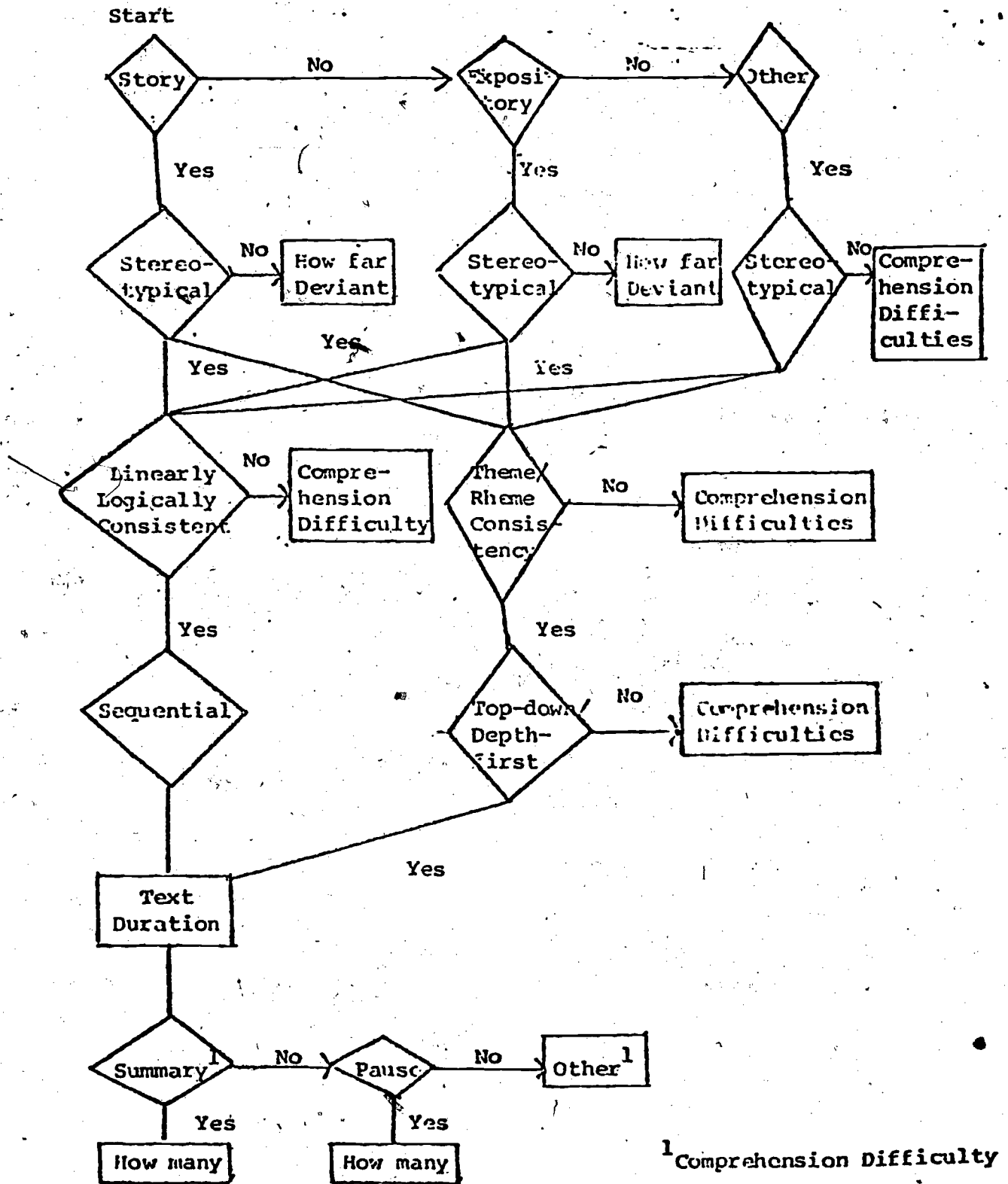
Text order, text structure and logical order were also decisive in distinguishing between the hard and easy text group. If the structure of a text is known by the reader, and as such matches a frame present in the reader, then such a text is easier to process than if this is not the case. Linear, sequential and logical order are also contributive to comprehension. But it has also become clear that it is not because a text belongs to a certain type of text (e.g., story vs. essay) that it would be easier to process. It is rather the intricate net of characteristics per text and the reader's aptitude for processing that specific text which is decisive.

The analysis of text structure may be summarized in a flowchart-like conclusion, in which the relation between the narrational analysis and text difficulty is shown (Figure 1).

4.2.2. Propositional Analysis

The objective of the propositional analysis was to bring together a more psycho-linguistically oriented analysis of language organization with present theory concerning the reader's processing activity. On this basis, a model of the parameters of discourse difficulty could be developed. For this model of representation of meaning, we relied upon Kintsch (1977) for the micro-structural analysis, and on Kintsch and Van Dijk (1976) and Van Dijk (1977) for the macro-structural analysis.

Figure 1. Flowchart of Ease of Comprehension by Means of Structure.



In the formation of the propositional structures four main determinants of processing difficulty are central:

- a. the number of propositions;
- b. the number of different arguments in relation to the coherent network of:
 - the number of unconnected graphs (propositions which share an argument are connected in a graph),
 - the number of items of given and new information,
 - the number of items of foregrounded information (i.e.; psychologically salient but linguistically unstressed information), ^{p 114}
 - the number of reinstatements (the function of a reinstatement is to link propositions in subgraphs and in between graphs in order to make a coherent text base),
 - the number of reorganizations (the function of a reorganization is to link propositions which belong together on a higher level).

We hypothesized that the more coherent a text is, and the more it is explicit about its macro-structures, the better the processing of the macro-structures.

The results of the propositional analysis can be read from Table 1. These results can be explained and related to the previous parameters in the light of "lexical density" (Halliday, 1980). "Lexical density" means that there are a large number of lexical items (including quite difficult words of fairly low frequency) packed into a rather simple grammatical structure, making written language denser than spoken language (Halliday, 1979). The more lexically dense a text is, the harder it may be to process the text according to the results of Table 1.

By means of this "highest common factor" (lexical density) the propositional analysis can be linked to the previously found parameters on a micro-structure level.

a. Syntax

It is not sentence length or sentence complexity of written language which makes processing harder, but rather how the syntax creates lexical density.

b. Lexicon

A more concrete and familiar lexicon, which is also used as a defining vocabulary, reduces the processing difficulty of lexically dense texts.

c. Information structure

It is not solely the number of arguments which contribute highly to the lexical density of texts, but whether the information presented is presumed (by the writer) to be known (by the reader) or not, and how that is reflected in the text. Here lies the importance of the distribution of "given" and "new," "defined" and "undefined" information.

In short, propositional analysis provides a theoretical basis for discourse processing. A list of parameters is therefore justified as a summary of the propositional and linguistic analysis (Table 1). On the basis of the list, compiled on a quantified basis, a theoretical model was formed (Figure 2).

5. Testing of the Model of Text Difficulty

The previous model gives rise to a number of hypotheses which were tested using a Pearson Product-Moment Correlation Matrix to investigate

- a. which of the nineteen variables correlated at a significant level with the dependent variable, comprehension difficulty;
- b. which variables correlate highly with each other so that they can be treated as a group;
- c. which of the variables both correlate with comprehension difficulty and with each other, so as to constitute clear clusters (Table 2).

The following hypotheses were tested:

General Hypothesis

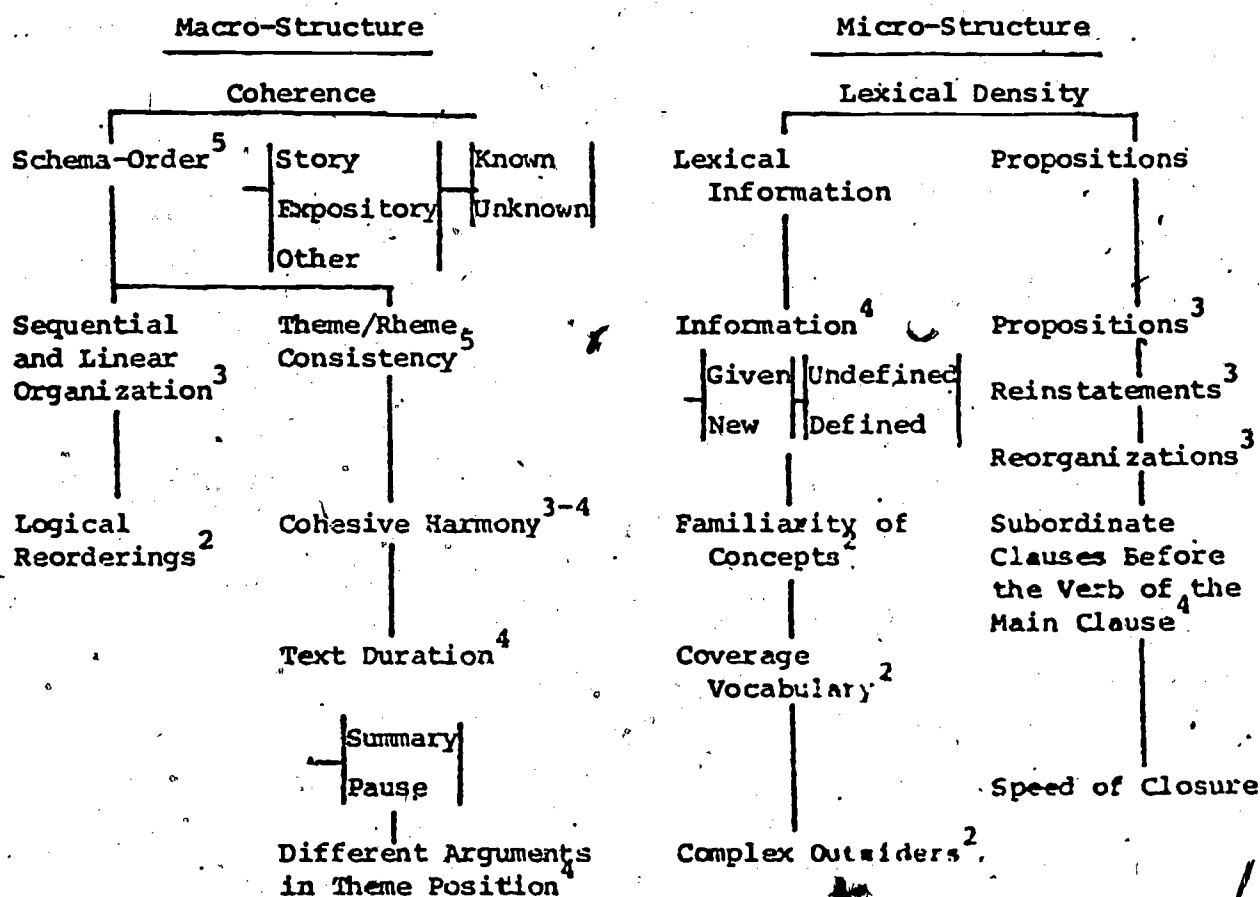
Macro-structural and micro-structural parameters both influence comprehension. Higher-order parameters in the model do not necessarily have more impact on comprehension. We expect that the parameters are inter-related and they can influence comprehension to a varying degree.

This hypothesis is confirmed. We see that clear clusters are formed by:

- a. the parameters of narrative condition, theme-rheme consistency, cohesive harmony and sequential and linear organization, all of which are macro-structural parameters, high up in the model, and
- b. the following parameters on the micro-structural level, at a higher or lower level: familiarity, coverage vocabulary, summary and number of propositions.

These parameters intercorrelate and also correlate with comprehension. Therefore we can conclude that in the texts we have been studying as they relate to the scores of the students, these eight parameters are the most significant ones.

Figure 2. A Model of the Parameters Indicating the Degree of Difficulty of Discourse Processing.



¹ In relation to 'Common ground' (Clark and Marshall, 1981).

² Computerized analysis is advisable.

³ Thorough manipulations are necessary (manual).

⁴ Fairly fast, but manipulations are necessary.

⁵ Fast, direct, no manipulations necessary (if some experience is present).

Table 16. Matrix of Intercorrelations Between Variables X_{1-19} and Y .

PAIRS OF INTERCORRELATIONS BETWEEN VARIABLES X_{1-19} AND Y	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}	X_{16}	X_{17}	X_{18}	X_{19}	Y
COMPLEX OUTSIDERS X_1	.25	.26	.63 ¹	-.55	.30	.24	-.55	-.20	.32	-.10	.35	-.66	.71	0.0	.64	0	0.0	0.0	.00
COVERAGE VOCABULARY X_2		.55	.20	.17	.91 ¹	.19	.10	.64 ²	.63	.17	.69 ²	.20	.25	.24 ²	.50	.76 ²	.74 ²	.74 ²	.66 ¹
FAMILIARITY X_3			-.20	.29	.32	-.21	-.54	.75 ²	.70 ²	-.25	.75 ²	-.32	-.35	.94 ¹	.22	.94 ¹	.94 ¹	.94 ¹	.69 ²
SUBORDINATE CLAUSE BEFORE CLOSE X_4				-.65 ²	.55 ²	.35	.64 ²	-.20	.20	.20	.20	.29	.60 ²	-.62	.51	-.63	-.63	-.63	.19
SPICE OF CLOSURE X_5					-.11	-.15	-.34	.14	0.0	-.26	-.05	-.10	-.58	.24	-.60	.24	.24	.24	-.66
PROPOSITIONS X_6						.66	.66	.52	.45	.35	.50 ²	.40	.52	.52	.64 ²	.53	.53	.53	.77 ²
REINSTATEMENTS X_7							.70 ¹	.17	.60	.53	.19	.17 ²	.62 ²	-.69	.62 ²	-.69	-.69	-.69	.67
REORGANIZATIONS X_8								-.10	.61	.46	-.61	.39	.70 ²	-.33	.50 ²	-.33	-.33	-.33	-.16
GIVEN-DEFINED ² X_9									.51	.16	.67 ²	0.0	-.61	.73 ²	.45	.73 ²	.73 ²	.73 ²	.76 ²
NEW-DEFINED X_{10}										-.60	.75 ²	-.33	.16	.71 ²	.50	.71 ²	.71 ²	.71 ²	.61
GIVEN-UNDEFINED X_{11}											.11 ²	.24	.15	-.66	.20	-.66	-.66	-.66	.10
SUPPLANT X_{12}												-.60	.10	.67 ¹	.74 ²	.67 ¹	.67 ¹	.67 ¹	.67 ²
PARSE X_{13}													.33	-.27	-.67	-.27	.27	-.27	.20
RECORDING X_{14}														-.20	.74 ²	-.20	.20	-.20	.69
NARRATIVE CONDI- TION X_{15}															.60	1.00	1.00	1.00	.70 ¹
DIFFERENT ARGU- MENTS IN THEM X_{16}																.60	.60	.60	.60
THREE-WINE CON- SISTENCY X_{17}																	1.00	1.00	.70 ¹
COHESIVE NOUNS AND LEXICAL CHAINS X_{18}																		1.00	.70 ¹
SEQUENTIAL AND LINEAR ORGA- NIZATION X_{19}																			.70 ¹
1 = significant at the .01 level																			
2 = significant at the .05 level																			

Hypothesis A

Parameters indicating ease of comprehension are: narrative condition, theme-rheme consistency, sequential and linear order, cohesive harmony, summary (on the macro-structural level); familiarity, coverage vocabulary, given and defined concepts, new and defined concepts and speed of closure (on the micro-structural level).

This hypothesis was also confirmed, as the intercorrelations between these parameters are very high (significant at the .05 and .01 level).

Hypothesis B

Comprehension difficulty is indicated by the number of complex outsiders, by the number of different arguments in theme position on the lexical level, by the number of subordinate clauses to the left of the finite verb of the main clause, by the number of reorganizations, reinstatements, logical reorderings and pauses on the syntactic and structural levels.

This hypothesis is also confirmed because the intercorrelation between these parameters is very significant as is their correlation with the dependent variable (significant at the .05 and .01 level). They correlate negatively at a significant level with the parameters indicating ease of comprehension.

Sub-Hypothesis 1

The number of propositions in itself is not an indicator of ease of difficulty in comprehension, but its relation to other parameters indicating ease or difficulty is important.

This hypothesis is also confirmed. The results indicate that the parameter "number of propositions" correlates both with comprehension score and with other parameters. It correlates with the parameters indicating difficulty (number of different arguments in theme position) and it correlates negatively with a parameter indicating ease of comprehension (speed of closure). It only correlates positively with one other parameter of comprehension ease, coverage vocabulary. A possible explanation for this may be that in the easy expository texts we have studied, more coverage vocabulary was used than in the hard texts. This would confirm that the number of propositions alone is not decisive either for easy or hard expository or non-expository texts.

Sub-Hypothesis 2

Theme-rheme consistency is one of the basic parameters on the text structural level because it relates to both macro-propositional and micro-propositional parameters.

It is not true that the theme-rheme parameter directly relates to the micro-propositional parameters. It does not correlate with the number of subordinate clauses, nor does it correlate negatively at a significant level with the number of different arguments used in theme position. But it does correlate significantly with other parameters of ease of comprehension (particularly given and defined information) and it correlates negatively with parameters indicating comprehension difficulty.

Sub-Hypothesis 3

The number of given-defined concepts is a significant parameter of ease of comprehension.

This hypothesis is confirmed. It correlates with comprehension at a .05 level of significance. Furthermore, it correlates with other parameters of ease of comprehension. This would point at the fact that "common ground" or "shared knowledge" is an important factor of comprehension. We have not studied this in depth, however, as such a study would entail a detailed cultural and knowledge analysis of the readers involved.

6. Discussion

As this study concerned reading in mother tongue and as the IEA tests were administered in various countries, we had various languages at our disposal. Therefore, the model was applied to other languages as well, in order to test its transferability. We want to stress that we selected languages (English, French and Dutch) related both in terms of the language organization and to the shared "knowledge" or "common ground" of the readers. The same method of determining the parameters was applied. No adaptations as to the method or to the parameters were necessary; yet, insights as to information structure, theme-rheme distribution, and cohesive harmony were strongly confirmed in both translations.

From the applications to other languages, we concluded that reading strategies which could evolve from the model are cognitively based and that, as such, they are transferable to the foreign language learning and teaching situation. But, rather than drawing conclusions at this level, we suggested that research in the foreign language should take up the thread and investigate readability in a foreign language at this cognitive level.

Implications for Comprehension and Readability Research .

- a. It has been shown that comprehension takes place on both micro- and macro-levels. Thus, learning to "construct" sentences is not enough in the comprehension process. But, on the other hand, the over-all format (genre and structure) is not the sole key either. For example, a story may be easy, but not for the mere fact that it is a story. All in all, the lexicon, on the micro-level and on the macro-level (i.e., lexical chains) plays a very important role in comprehension.
- b. The study has also shown that both reader and text contribute to the construction of meaning. Language in its functional and structural organization should be given due attention in reading and writing. This means that:
 - It is not possible to write a readable text for a certain age level solely on the basis of a cook book recipe. The cook has to know his guests quite well, because "fast food" is not easily digested by them. We suggest that the parameters indicating the textual characteristics should be carefully studied in close relation to the potential readers in order to write a text which facilitates its processing.
 - The active contribution of the reader in constructing meaning should be focused upon and exploited. As constructing meaning is a cognitive activity, it is possible to put the tools in the hands of readers. Thus, they can actively direct their comprehension by being provided with effective strategies.
- c. Here lies the opportunity for the teacher to assist the reader in this process. "Process" is the right word, because reading strategies are necessary here. The teacher cannot only focus on the product, the test scores, but has to concentrate on the process leading to the outcome and analyze the outcome from the point of view of the processing by the reader of the specific text(s).

Thus, the input is the text, the output is the score. The "go-between" is the reader, who, if helped by the teacher and the right reading strategies, can learn not only to be the recipient of the message but to actively and cognitively interact with it and as such make meaning. Such a strategy puts the reader on the road to self-instruction and self-development. And is this not what effective reading is all about?

Finally, the study has shown that comprehension is not just a set of mental processes which can be defined independently of language. Rather, it is a set of processes which operate on specific features of language. But on this view, the reader himself takes a central position. He is the one who makes the text into a text. This study has not attempted to define HOW a potential reader utilizes his prior knowledge or previous experience when he comprehends or judges from a text. This also remains a main difficulty, which needs further research.

REFERENCES

- Anderson, R. C., Spiro, R. J., and Montague, W. E. (Eds.), Schooling and the Acquisition of Knowledge, Hillsdale, N. J.: LEA, 1977.
- Backman, J., "Reading Comprehension and Perceived Comprehensibility of Lexical Density at Discourse and Sentence Level", Educational Reports Umeå No. 10, Umeå University and School of Education, Sweden, 1976.
- Bartlett, J. R., Remembering, Cambridge: Cambridge University Press, 1932.
- Bormuth, J. R., "Readability: A New Approach", Reading Research Quarterly, 3, 1, 1966, pp. 81-132.
- Bormuth, J. R., Development of Readability Analyses, U.S. Office of Education, Project No. 7-0052, Chicago: University of Chicago, 1969.
- Bormuth, J. R., Development of Standards of Readability: Toward a Rational Criterion of Passage Performance, Report No. 9-0237, Chicago: University of Chicago, 1971.
- Bormuth, J. R., "An Operational Definition of Comprehension Instruction", in Goodman, K. S. and Fleming, J. T. (Eds.), Psycholinguistics and the Teaching of Reading, Newark, Del.: IRA, 1977, pp. 48-61.
- Bransford, J. D., Barclay, J. R., and Franks, J. F., "Sentence Memory: A Constructive Versus Interpretive Approach", Cognitive Psychology, 3, 1972, pp. 193-209.
- Carroll, J. B., "Defining Language Comprehension: Some Speculations", in Freedle, R. O., and Carroll, J. B. (Eds.), Language Comprehension and the Acquisition of Knowledge, New York: Winston and Wiley, pp. 11-31.
- Carroll, J. B., "Developmental Parameters of Reading Comprehension", in Guthrie, J. T. (Ed.), Cognition, Curriculum and Comprehension, Newark, Del.: IRA, 1977, pp. 1-15.
- Chafe, W. L., "Meaning and the Structure of Language", Language, 48, 1972a, pp. 134-160.
- Chafe, W. L., "Discourse Structure and Human Knowledge", in Freedle, R. O., and Carroll, J. B. (Eds.), Language Comprehension and the Acquisition of Knowledge, New York: Winston and Wiley, 1972b, pp. 41-71.
- Chafe, W. L., "Givenness, Contrastiveness, Definiteness, Subjects, Topics, and Point of View", in Li, C. N. (Ed.), Subject and Topic, New York: Academic Press, 1974, pp. 27-55.
- Chomsky, N., Syntactic Structures, The Hague: Mouton, 1957.

Christensen, F., "Symposium on the Paragraph", in College Composition and Communication, 17, 2, 1966, pp. 60-67.

Clark, E., and Clark, H. H., Psychology and Language: An Introduction to Psycholinguistics, New York: Harcourt Brace Jovanovich, 1977.

Clark, H. H., and Carlson, T. B., "Context for Comprehension", in Long, J., and Baddeley, A. (Eds.), Attention and Performance IX, Hillsdale, N. J.: LEA, 1981.

Clark, H. H., and Marshall, C. E., "Definite Reference and Mutual Knowledge", in Joshi, A. K., and Webber, B. (Eds.), Elements in Discourse Understanding, Cambridge: Cambridge University Press, 1981.

Cófer, C. N., "A Comparison of Logical and Verbatim Learning of Prose Passages of Different Lengths", in The American Journal of Psychology, 54, 1941, pp. 1-20.

Davison, A., Kantor, R. N., Hannah, J., Hermon, G., Lutz, R., and Salzillo, R., Limitations of Readability Formulas in Guiding Adaptations of Texts, Technical Report No. 162, Center for the Study of Reading, University of Illinois, 1980.

Daneš, F., "Functional Sentence Perspective and the Organization of the Text", in Daneš, F. (Ed.), Papers on Functional Sentence Perspective, The Hague: Mouton, 1974, pp. 106-128.

Engels, L. K., "The Fallacy of Word Counts", in IRAL 6, 3, 1968, pp. 213-231.

Engels, L. K., "Pedagogical Grammars", Paper No. 20, Series B, Trier: LAUT, 1977.

Engels, L. K., "Pedagogical Grammar as Applied Semiotics", in Matter, F. (Ed.), Taaltoetsen in Onderwijs en Onderzoek, Amsterdam VU en Diepenbeek Wetenschappelijk Onderwijs Limburg, 1978, pp. 109-122.

Engels, L. K., "Testing and Mastery Learning of English Vocabulary at University Level", Paper given at ANELA, Tilburg, (forthcoming).

Engels, L. K., Geens, D., and Martin, W., "Analysis of Present-Day English Theatrical Language, 1966-72", PAL, 6, 1975.

Engels, L. K., Van Beckhoven, B., Leenders, T., and Brasseur, I., L.E.T. (Leuven English Teaching) Vocabulary List Based on Objective Frequency Combined with Subjective Word Selection, Leuven, Belgium: Catholic University Leuven, Acco, 1981.

Fishman, A. S., Elements of Cohesion: the Effect of Noun Phrase Organizers and Anaphoric References on Paragraph Comprehension, Unpublished Ph.D. Dissertation, Northwestern University, Evanston, Illinois, 1977.

Flesch, R. F., "A Readability Yardstick", Journal of Applied Psychology, 32, 1948, pp. 221-233.

- Fløttum, K., "Vanskelighetsgrad ved Norske Lesestykker for Utlendinger", University of Trondheim, 1980. (Unpublished paper)
- Fries, P. H., "On the Status of Theme in English: Arguments from Discourse", Central Michigan University, 1980. (Unpublished paper)
- Gagné, R., and Wiegand, V. K., "Effects of a Superordinate Context on Learning and Retention of Facts", Journal of Educational Psychology, 61, 1970, pp. 408-409.
- Genette, G., Figures, III, Discours du Récit, Paris: Sévil, 1972.
- Goodman, K. S., "Linguistically Sound Research in Reading", in Farr, R., Weintraub, S., and Tone, B. (Eds.), Improving Reading Research, Newark, Del.: IRA, 1977.
- Halliday, M. A. K., Intonation and Grammar in British English, The Hague: Mouton, 1967.
- Halliday, M. A. K., "Notes on Transitivity and Theme in English, Parts 1-3", Journal of Linguistics, 3, 1; 3, 2, 1967, pp. 199-244; 4, 2, pp. 179-215.
- Halliday, M. A. K., "Language Structure and Language Function", in Lyons, J. (Ed.), New Horizons in Linguistics, Baltimore: Penguin Books, 1970a, pp. 140-165.
- Halliday, M. A. K., "Functional Diversity in Language as Seen From a Consideration of Modality and Mood in English", Foundations of Language, 6, 3, 1970b, pp. 322-361.
- Halliday, M. A. K., Explorations in the Functions of Language, London: Edward Arnold, 1977.
- Halliday, M. A. K., Language as a Social Semiotic: the Social Interpretation of Language and Meaning, London: Edward Arnold, 1978.
- Halliday, M. A. K., "Differences Between Spoken and Written Language: Some Implications for Literacy Teaching", in Page, G., Elkins, J., and O'Connor, B. (Eds.), Communication Through Reading: Proceedings of the Fourth Australian Reading Conference, Vol. 2, Adelaide: Australian Reading Association, 1979, pp. 37-52.
- Halliday, M. A. K., "How is a Text Like a Clause?" Paper presented at the Nobel Symposium on Text Processing, Stockholm, 1980.
- Halliday, M. A. K., Introduction in Functional Grammar: Analysis of Texts in Linguistics on Clause Level and Beyond, London: Edward Arnold, (forthcoming).
- Halliday, M. A. K., "Text Semantics and Clause Grammars: Some Patterns of Realization", (forthcoming).

- Halliday, M. A. K., and Hasan, R., Cohesion in English, London: Longman, 1976.
- Halliday, M. A. K., and Hasan, R., Text and Context: Aspects of Language in a Social-Semiotic Perspective, (in press).
- Hasan, R., "Linguistics and the Study of Literary Texts"; Etudes de Linguistique Appliquée, 5, 1967, pp. 106-121.
- Hasan, R., "Grammatical Cohesion in Spoken and Written English, Part I", Program in Linguistics and English Teaching, Paper No. 7, London: Longmans, 1968.
- Hasan, R., "Rime and Reason in Literature", Chatman, S. (Ed.), Literary Style: A Symposium, London: Oxford University Press, 1971.
- Hasan, R., "Coherence and Cohesive Harmony", (forthcoming).
- Hausenblas, K., "On the Characterization and Classification of Discourse", Travaux Linguistique de Prague, 1964, p. 1.
- Haviland, S. E., and Clark, H. H., "What's New? Acquiring New Information as a Process in Comprehension", Journal of Verbal Learning and Verbal Behaviour, 3, 1974, pp. 512-521.
- Hunt, K. W., Grammatical Structures Written at Three Grade Levels, Report No. 3, National Council of Teachers of English, Urbana, Ill.: NCTE, 1965.
- Ingram, E., "Vanskelighetsgrad ved norske tekster for utlendinger: oppfølging og validitetstesting". (Unpublished paper)
- Johanson, S., "Report from a Symposium on Grammatical Tagging of English Text Corpora", Association for Literary and Linguistic Computing Bulletin, 7, 3, 1979, pp. 301-302.
- Just, M. A., and Carpenter, P. A. (Eds.), Cognitive Processes in Comprehension, Hillsdale, N. J.: LEA, 1977.
- Kieras, D. E., "Good and Bad Structure in Simple Paragraphs: Effects on Apparent Theme, Reading and Time Recall", Journal of Verbal Learning and Verbal Behavior, 17, 1978, pp. 13-28.
- Kintsch, W., "On Comprehending Stories", in Carpenter, M. A., and Just, P. A. (Eds.), Cognitive Processes in Comprehension, Hillsdale, N. J.: LEA, 1977, pp. 33-62.
- Kintsch, W., and Keenan, J. M., "Comprehension and Recall of Text as a Function of Content Variables", in Journal of Verbal Learning and Verbal Behavior, 14, 1975, pp. 196-214.
- Kintsch, W., and Van Dijk, T. A., "Cognitive Psychology and Discourse Recalling and Summarizing Stories", in Dressler, W. U. (Ed.), Trends in Test-Linguistics, New York/Berlin: de Gruyter, 1976.

- Kintsch, W., and Vipond, D., "Reading Comprehension and Readability in Educational Practice and Psychological Theory", in Nilsson, L. G. (Ed.), Memory and Cognition, Hillsdale, N. J.: LEA, 1979.
- Klare, G., "Assessing Readability", Reading Research Quarterly, 10, 1, 1974, pp. 64-102.
- Klare, G., Shuford, E., and Nichols, W., "The Relation of Style Difficulty, Practice and Ability to Efficiency of Reading and to Retention", Journal of Applied Psychology, 41, 1957, pp. 222-226.
- Kucera, H., and Francis, W. N., Computational Analysis of Present-Day American English, Providence, R. I.: Brown University Press, 1967.
- Longman's Dictionary of Contemporary English, 1st Ed., London: Longman, 1978.
- Mac Whinney, B., and Bates, E., "Sentential Devices for Conveying Givenness and Developmental Study", Journal of Verbal Learning and Verbal Behavior, 17, 1978, pp. 539-558.
- Mathesius, V., "O tak Zvanem aktualnim clemení Vetnem", Slovo a slovesnost, 5, 1939, pp. 171-174.
- Meyer, B., The Organization of Prose and its Effects on Memory, Amsterdam: North Holland Publications, 1975.
- Meyer, B., "The Structure of Prose", in Anderson, R. G., and Spiro, R. J. (Eds.), Schooling and the Acquisition of Knowledge, 1977, pp. 179-201.
- Miller, G. A., Language and Communication, New York: McGraw-Hill, 1951.
- Miller, G. A., "The Magic Number 7, Plus or Minus 2", Psychology of Communication, New York, 1967.
- Minsky, M., "A Framework for Representing Knowledge", in Winston, P. (Ed.), Psychology of Computer Vision, New York: McGraw-Hill, 1975, pp. 211-278.
- Mountford, A., Discourse Analysis and the Simplification of Reading Material, Unpublished M.Litt. Thesis, University of Edinburgh, Scotland, 1975.
- Osgood, E., "Where do Sentences Come From?", in Steinberg, D. D., and Jakobovits, L. A. (Eds.), Semantics, Cambridge: Cambridge University Press, 1971.
- Pearson, P. D., "The Effects of Grammatical Complexity on Children's Comprehension and Conception of Certain Semantic Relations", Reading Research Quarterly, 10, 2, 1973, pp. 155-191.
- Pearson, P. D., and Johnson, D. D., Teaching Reading Comprehension, New York: Holt, Rinehart, Winston, 1978.

- Peel, E. A., "The Analysis of Comprehension and Judgement from Textual Material", Educational Review, 27, 2, 1975.
- Perfetti, C. A., "Language Comprehension and Fast Decoding: Some Psycholinguistic Prerequisites for Skilled Reading Comprehension", in Guthrie, J. T. (Ed.), Cognition, Curriculum and Comprehension, 1977a, pp. 20-41.
- Perfetti, C. A., "Comments on Five Exemplary Reading Programs", in Guthrie, J. T. (Ed.), Cognition, Curriculum and Comprehension, 1977b, pp. 179-183.
- Perfetti, C. A., and Lesgold, A. M., "Discourse Comprehension and Sources of Individual Differences", in Carpenter, M. A., and Just, P. A., Cognitive Processes in Comprehension, 1977, pp. 141-184.
- Pools, P., Robberechts, K., and Van de Putte, H., A 2000-Word Frequency List Based on the Brown University Corpus and the Theatre Corpus, Unpublished M.A. Thesis, University of Leuven, 1978.
- Purves, A. C., Literature Education in Ten Countries, an Empirical Study, Stockholm: Almqvist and Wiksell, 1973.
- Purves, A. C., "Putting Readers in Their Places: Some Alternatives to Cloning Stanley Fish", College English, National Council of Teachers of English, 1980, pp. 228-236.
- Purves, A. C., Harnisch, D. L., Bauer, B., and Quirk, D. L., Reading and Literature: American Achievement in International Perspective, Research Report No. 20, National Council of Teachers of English, Urbana, Ill.: NCTE, 1981.
- Quirk, R., Greenbaum, S., and Svartvik, F., A Grammar of Contemporary English, London: Longman, 1972.
- Richards, J. C., "Word Lists: Problems and Prospects", RELJ Journal, 5, 2, 1974, pp. 1-18.
- Scharer, G. D. E., The Structure and Paragraphing of Expository Prose with Reference to the Writing Problems of the Advanced Second Language Learner, Unpublished Ph.D. Dissertation, University of Edinburgh, 1980.
- Schlessinger, I. M., Sentence Structure and the Reading Process, The Hague: Mouton, 1968.
- Simensen, A. M., "Vanskelighetsgrad i Norske Lesetekster", University of Trondheim, 1980. (Unpublished paper)
- Thorndike, R., Reading Comprehension Education in 15 Countries, Stockholm: Almqvist and Wiksell, 1973.

- Thorndyke, P. W., "Inferences in Discourse Comprehension", Journal of Verbal Learning and Verbal Behavior, 15, 4, 1976, pp. 437-446.
- Thorndyke, P. W., "Cognitive Structures in Comprehension and Memory of Narrative Discourses", Journal of Verbal Learning and Verbal Behavior, 16, 1, 1977, pp. 77-109.
- Thurstone, L. L., "A Factorial Study of Perception", Psychometric Monographs, No. 4, Chicago: University of Chicago Press, 1944.
- Thurstone, L. L., and Jeffrey, T. E., "Closure Speed", Test Administration Manual, Revised, Chicago: University of Chicago Press, 1966.
- Van Dijk, T. A., "Models of Macro-Structures", Preliminary Working Draft, University of Amsterdam, 1974.
- Van Dijk, T. A., "Recalling and Summarizing Complex Discourse", Preliminary Working Draft, University of Amsterdam, 1975a.
- Van Dijk, T. A., "Discourse Meaning and Memory", Journal of Reading Behavior, 1975b.
- Van Dijk, T. A., "Semantic Macro-Structures and Knowledge Frames in Discourse Processing", in Carpenter, M. A., and Just, P. A. (Eds.), Cognitive Processes in Comprehension, Hillsdale, N. J.: LEA, 1977, pp. 3-32.
- Van Dijk, T. A., Text and Context: Explorations in the Semantics and Pragmatics of Discourse, London: Longman, 1980.
- Van Hauwermeiren, P., Het Leesbaarheidsonderzoek (Met Toepassing op Nederlandse Teksten), Unpublished Ph.D. Dissertation, Catholic University Leuven, 1972.
- Van Hauwermeiren, P., "Leesbaarheidspredictie?", Instituut voor Vertalers en Tolken, 1975, pp. 10-25.
- Van Parreren, C. F., Psychologie van het Leren II, Actualisering van Leerresultaten, Arnhem: Van Loophum Slaterus, 1970.
- Wardrop, J. L., "Design Problems in Reading Research", in Farr, R., Weintraub, S., and Tone, B. (Eds.), Improving Reading Research, Newark, Del.: IRA, 1977.
- West, M., A General Service List of English Words, with Semantic Frequencies and a Supplementary Word List for the Writing of Popular Science and Technology, 10th Ed., London: Longman, 1977.
- Widdowson, H. G., and Urquhart, A. H., K.A.A.U. English for Academic Purposes Report, University of Edinburgh, 1976. (Unpublished).

Widdowson, H. G., "The Realization of Rules in Written Discourse", Recherches et échanges, 4, 2, 1979, pp. 1-19.

Winograd, T., "A Framework for Understanding Discourse", in Carpenter, M. A., and Just, P. A. (Eds.), Cognitive Processes in Comprehension, Hillsdale, N. J.: LEA, 1977, pp. 63-88.